

Weighted finite automata with output

Jelena Ignjatović · Miroslav Ćirić · Zorana Jančić

Received: date / Accepted: date

Abstract In this paper we prove the equivalence of sequential, Mealy-type and Moore-type weighted finite automata with output, with respect to various semantics which are defined here.

Keywords Weighted automaton · Fuzzy automaton · Sequential automaton · Mealy-type automaton · Moore-type automaton

1 Introduction

Finite automata with output are a simple mathematical model of computation with numerous applications in different areas. Generally speaking, the main role of an automaton with output is to transform finite sequences of input symbols to finite sequences of output symbols, and its behavior is understood as a function or relation between the sets of all input and output sequences. The most simple among such automata, the ordinary deterministic finite automata with output, have two basic models. Mealy-type automata simultaneously pass into a new state and emit output, and the value of the output depends both on the current state and the current input, whereas Moore-type automata emit output just after

the transition to the next state, and the value of the output depends solely on this new state. Although different, these two models are equivalent, in the sense that any Mealy-type automaton can be converted to a Moore-type automaton with the same behavior, and vice versa.

When dealing with more complex types of automata, such as, for instance, fuzzy or weighted finite automata with output, things become more complicated. Fuzzy finite automata with output have been studied by many authors who have considered several different models and semantics. Sequential fuzzy finite automata, where both transitions and outputs are modeled by a single transition-output function, have been investigated in [3,6,7,8,10,13]. It should be noted that a sequential fuzzy finite automaton with an input alphabet X and an output alphabet Y can be considered as a fuzzy automaton (i.e., fuzzy transition system) with the input alphabet $X \times Y$ and without output. On the other hand, the articles [2,3,4,11,12] have dealt with Mealy-type and Moore-type fuzzy finite automata, where transitions and outputs are modeled by separate transition and output functions, and the behavior of these automata has been defined in different ways.

All the mentioned models of automata are defined here in a more general context, for weighted finite automata over a semiring. Besides, the behavior of Mealy-type weighted finite automata is defined in three different ways – we distinguish the $n1$ -semantics, the $1n$ -semantics and the sequential semantics, whereas for Moore-type automata we distinguish the $n1$ -semantics and the $1n$ -semantics. In the framework of Mealy-type and Moore-type fuzzy finite automata the $n1$ -semantics has been considered in [1,2,4,12], the $n1$ -semantics in [3,11], and the sequential semantics in [3]. The purpose of the paper is to study the equivalence between the mentioned types of weighted finite automata with out-

Research supported by Ministry of Education, Science and Technological Development, Republic of Serbia, Grant No. 174013.

J. Ignjatović · M. Ćirić · Z. Jančić
University of Niš, Faculty of Sciences and Mathematics,
Department of Mathematics and Computer Science
Višegradska 33, 18000 Niš, Serbia
Tel.: +38118224492
Fax: +38118533014
E-mail: jelena.ignjatovic@pmf.edu.rs, miroslav.ciric@pmf.edu.rs,
zorana.jancic329@gmail.com

puts, with respect to the mentioned semantics. We show that each Mealy-type weighted finite automaton can be converted into a sequential weighted finite automaton equivalent w.r.t. the sequential semantics, each Moore-type weighted finite automaton can be converted into a sequential weighted finite automaton equivalent w.r.t. the $1n$ -semantics, and vice versa, every Mealy-type weighted finite automaton can be converted to a Moore-type weighted finite automaton equivalent w.r.t. both the $1n$ -semantics and $n1$ -semantics, and each Moore-type weighted finite automaton can be converted to a Mealy-type weighted finite automaton equivalent w.r.t. the $1n$ -semantics. Moreover, we determine certain conditions under which a sequential weighted finite automaton can be converted to a Mealy-type weighted finite automaton equivalent w.r.t. the sequential semantics. In all these cases we also estimate the growth of the number of states during the conversion.

Note that although different models of fuzzy automata with output were studied in numerous papers, only the paper of Li and Pedrycz [3] discussed the equivalence of these models, and our work is a continuation of this research.

The paper is organized as follows. In Section 2 we recall basic notions and notation concerning semirings and matrices over a semiring, and in Section 3 we present definitions of sequential weighted automata and their behavior. Thereafter, in Sections 4 and 5 we define Mealy-type and Moore-type weighted finite automata, three different semantics for Mealy-type weighted automata and two semantics for Moore-type weighted automata. Our main results are presented in Section 6, where we prove the equivalence of sequential, Mealy-type and Moore-type weighted finite automata with respect to various semantics. Finally, in Section 7 we consider crisp-deterministic Mealy-type and Moore-type weighted finite automata and show that all previously considered semantics coincide for such automata.

2 Preliminaries

Throughout this paper, \mathbb{N} denotes the set of natural numbers (without zero), X^+ and X^* denote respectively the free semigroup and the free monoid over an alphabet X , and ε denotes the *empty word* in X^* .

A *semiring* is a structure $(S, +, \cdot, 0, 1)$ consisting of a set S , two binary operations $+$ and \cdot on S , and two constants $0, 1 \in S$ such that the following is true:

- (i) $(S, +, 0)$ is a commutative monoid,
- (ii) $(S, \cdot, 1)$ is a monoid,
- (iii) the distributivity laws $(r + s) \cdot t = r \cdot t + s \cdot t$ and $t \cdot (r + s) = t \cdot r + t \cdot s$ hold for every $r, s, t \in S$,

- (iv) $0 \cdot s = s \cdot 0 = 0$ for every $s \in S$.

As usual, we identify the structure $(S, +, \cdot, 0, 1)$ with its carrier set S . A semiring S is called *additively idempotent* if $s + s = s$, for every $s \in S$, or equivalently, if $1 + 1 = 1$. For $n \in \mathbb{N}$ and $s \in S$, the n -th *additive power* of s is the element $ns = s + s + \dots + s$ (n times).

Let P and Q be sets. We let Q^P denote the set of all functions from P to Q . Next, let S be a semiring and let A be a finite non-empty set. A mapping $\mu : A \times A \rightarrow S$ is called an $A \times A$ -*matrix* over S , and a mapping $v : A \rightarrow S$ is called an A -*vector* over S . If S is a particular ordered set (e.g., the real unit interval $[0, 1]$), then matrices are called *fuzzy relations*, and vectors are called *fuzzy subsets* in the literature.

Given matrices $\mu_1, \mu_2 \in S^{A \times A}$ and vectors $v_1, v_2 \in S^A$. Then we define the *matrix product* $\mu_1 \cdot \mu_2 \in S^{A \times A}$, the *matrix-vector products* $v_1 \cdot \mu_1 \in S^A$ and $\mu_1 \cdot v_1 \in S^A$, and the *scalar product* $v_1 \cdot v_2 \in S$ as follows for every $a_1, a_2 \in A$:

$$\begin{aligned} (\mu_1 \cdot \mu_2)(a_1, a_2) &= \sum_{a \in A} \mu_1(a_1, a) \cdot \mu_2(a, a_2), \\ (v_1 \cdot \mu_1)(a_1) &= \sum_{a \in A} v_1(a) \cdot \mu_1(a, a_1), \\ (\mu_1 \cdot v_1)(a_1) &= \sum_{a \in A} \mu_1(a_1, a) \cdot v_1(a), \\ v_1 \cdot v_2 &= \sum_{a \in A} v_1(a) \cdot v_2(a). \end{aligned}$$

Recall that the addition of S is commutative and that A is non-empty; thus, the sums on the right-hand sides are well defined. Moreover, since distributivity of the multiplication operation over the addition operation holds, the matrix product and matrix-vector products are associative. The *Hadamard (pointwise) product* $v_1 \odot v_2$ of vectors $v_1, v_2 \in S^A$ is defined as follows for any $a \in A$:

$$(v_1 \odot v_2)(a) = v_1(a) \cdot v_2(a).$$

Given a vector $v \in S^A$, we define a matrix $D(v) \in S^{A \times A}$ as follows for every $a, b \in A$:

$$D(v)(a, b) = \begin{cases} v(a) & \text{if } a = b, \\ 0 & \text{otherwise.} \end{cases}$$

For an arbitrary matrix $\mu \in S^{A \times A}$ and $a, b \in A$, we can easily verify that

$$\begin{aligned} (D(v) \cdot \mu)(a, b) &= v(a) \cdot \mu(a, b), \\ (\mu \cdot D(v))(a, b) &= \mu(a, b) \cdot v(b). \end{aligned} \tag{1}$$

3 Sequential weighted automata

All weighted automata that will be discussed throughout this paper will have finite sets of states, input and output alphabets. Such automata are usually called weighted finite automata, but here we omit the adjective “finite” because it will entail.

A *sequential weighted automaton* over a semiring S is a tuple $\mathcal{A} = (A, X, Y, \sigma^A, \mu^A)$, where A, X and Y are finite non-empty sets, called respectively the *set of states*, the *input alphabet*, and the *output alphabet*, $\sigma^A : A \rightarrow S$ is the *initial weight vector* and $\mu^A : A \times X \times Y \times A \rightarrow S$ is the *weighted transition-output function*. The functions μ^A and σ^A can be understood as follows. When the weighted automaton \mathcal{A} is in a state $a \in A$ and it receives the input symbol $x \in X$, we can interpret $\mu^A(a, x, y, b)$ as the degree to which \mathcal{A} moves into a state $b \in A$ and emits the output symbol $y \in Y$. On the other hand, we can interpret $\sigma^A(a)$ as the degree to which $a \in A$ is an initial state. Without danger of confusion, in cases when we deal with a single sequential weighted automaton, we will omit the superscript A in σ^A and μ^A .

Let us note that the free monoid $(X \times Y)^*$ is isomorphic to the submonoid of $X^* \times Y^*$ consisting of all pairs $(u, v) \in X^* \times Y^*$ such that $|u| = |v|$, and we will identify these two monoids, as is commonly done in algebra. Thus, the identity in $(X \times Y)^*$ is identified with the identity $(\varepsilon, \varepsilon)$ of $X^* \times Y^*$, where ε denotes the identity (empty word) both in X^* and Y^* .

For any pair $(x, y) \in X \times Y$ we define $\mu_{x,y} : A \times A \rightarrow S$ by $\mu_{x,y}(a, b) = \mu(a, x, y, b)$, for all $a, b \in A$, and for any $(u, v) \in (X \times Y)^*$ the *weighted transition-output matrix* (or the *weighted transition-output relation*) $\mu_{u,v} : A \times A \rightarrow S$ is defined as follows: If $a, b \in A$, then

$$\mu_{\varepsilon, \varepsilon}(a, b) = \begin{cases} 1, & \text{if } a = b, \\ 0, & \text{otherwise,} \end{cases} \quad (2)$$

and if $a, b \in A$, $(u, v) \in (X \times Y)^+$ and $(x, y) \in X \times Y$, then

$$\mu_{ux,vy}(a, b) = \sum_{c \in A} \mu_{u,v}(a, c) \cdot \mu_{x,y}(c, b). \quad (3)$$

It is easy to check that

$$\mu_{up,vq}(a, b) = \sum_{c \in A} \mu_{u,v}(a, c) \cdot \mu_{p,q}(c, b), \quad (4)$$

i.e., $\mu_{up,vq} = \mu_{u,v} \cdot \mu_{p,q}$, for all $a, b \in A$ and $(u, v), (p, q) \in (X \times Y)^*$. Therefore, if $u = x_1 \dots x_n$ and $v = y_1 \dots y_n$, where $x_1, \dots, x_n \in X$ and $y_1, \dots, y_n \in Y$, then

$$\mu_{u,v}(a, b) = \sum_{(c_1, \dots, c_{n-1}) \in A^{n-1}} \mu_{x_1, y_1}(a, c_1) \cdot \mu_{x_2, y_2}(c_1, c_2) \cdot \dots \cdot \mu_{x_n, y_n}(c_{n-1}, b), \quad (5)$$

or, in the matrix form, $\mu_{u,v} = \mu_{x_1, y_1} \cdot \mu_{x_2, y_2} \cdot \dots \cdot \mu_{x_n, y_n}$.

Definition 3.1 The *behavior* of a sequential weighted automaton \mathcal{A} is the function $\llbracket \mathcal{A} \rrbracket : (X \times Y)^* \rightarrow S$ defined by

$$\llbracket \mathcal{A} \rrbracket(\varepsilon, \varepsilon) = \sum_{a, b \in A} \sigma(a) \cdot \mu_{\varepsilon, \varepsilon}(a, b) = \sum_{a \in A} \sigma(a), \quad (6)$$

and

$$\begin{aligned} \llbracket \mathcal{A} \rrbracket(u, v) &= \sum_{a, b \in A} \sigma(a) \cdot \mu_{u,v}(a, b) \\ &= \sum_{(a, a_1, \dots, a_n) \in A^{n+1}} \sigma(a) \cdot \mu_{x_1, y_1}(a, a_1) \cdot \mu_{x_2, y_2}(a_1, a_2) \cdot \dots \cdot \mu_{x_n, y_n}(a_{n-1}, a_n), \end{aligned} \quad (7)$$

for each $(u, v) \in (X \times Y)^+$, $u = x_1 \dots x_n$, $v = y_1 \dots y_n$, for some $n \in \mathbb{N}$, $x_1, \dots, x_n \in X$ and $y_1, \dots, y_n \in Y$. In other words,

$$\llbracket \mathcal{A} \rrbracket(u, v) = \sigma \cdot \mu_{u,v} \cdot \tau, \quad (8)$$

where $\tau : A \rightarrow S$ is given by $\tau(a) = 1$, for any $a \in A$.

4 Mealy-type weighed automata

A *Mealy-type weighted automaton* over a semiring S is a tuple $\mathcal{A} = (A, X, Y, \sigma^A, \delta^A, \omega^A)$, where A, X, Y and σ^A are as in the definition of a sequential weighted automaton, $\delta^A : A \times X \times A \rightarrow S$ is the *weighted transition function*, and $\omega^A : A \times X \times Y \rightarrow S$ is the *weighted output function*. The functions δ^A and ω^A can be understood as follows. When the automaton \mathcal{A} is in a state $a \in A$ and it receives the input symbol $x \in X$, we can interpret $\delta^A(a, x, b)$ as the degree to which \mathcal{A} moves into a state $b \in A$, and $\omega^A(a, x, y)$ as the degree to which \mathcal{A} emits the output symbol $y \in Y$. When we deal with a single Mealy-type weighted automaton, we omit the superscript A in σ^A , δ^A and ω^A .

For any $x \in X$ we define $\delta_x : A \times A \rightarrow S$ by $\delta_x(a, b) = \delta(a, x, b)$, for all $a, b \in A$, and for any $u \in X^*$ we define the *weighted transition matrix* (or *weighted transition relation*) $\delta_u : A \times A \rightarrow S$ as follows: For any $a, b \in A$ we set

$$\delta_\varepsilon(a, b) = \begin{cases} 1, & \text{if } a = b, \\ 0, & \text{otherwise,} \end{cases} \quad (9)$$

and if $a, b \in A$, $u \in X^*$ and $x \in X$, then

$$\delta_{ux}(a, b) = \sum_{c \in A} \delta_u(a, c) \cdot \delta_x(c, b). \quad (10)$$

It is easy to verify that

$$\delta_{uv}(a, b) = \sum_{c \in A} \delta_u(a, c) \cdot \delta_v(c, b), \quad (11)$$

for all $a, b \in A$ and $u, v \in X^*$, i.e., $\delta_{uv} = \delta_u \cdot \delta_v$. Hence, if $u = x_1 \cdots x_n$, for some $n \in \mathbb{N}$ and $x_1, \dots, x_n \in X$, then

$$\delta_u(a, b) = \sum_{(c_1, \dots, c_{n-1}) \in A^{n-1}} \delta_{x_1}(a, c_1) \cdot \delta_{x_2}(c_1, c_2) \cdot \dots \cdot \delta_{x_n}(c_{n-1}, b), \quad (12)$$

i.e., $\delta_u = \delta_{x_1} \cdot \delta_{x_2} \cdot \dots \cdot \delta_{x_n}$.

Next, define a vector $\omega_{\varepsilon, \varepsilon} : A \rightarrow S$ by $\omega_{\varepsilon, \varepsilon}(a) = 1$, for any $a \in A$, and for any pair $(x, y) \in X \times Y$ define a vector $\omega_{x, y} : A \rightarrow S$ by $\omega_{x, y}(a) = \omega(a, x, y)$, for every $a \in A$. For an arbitrary $(u, v) \in (X \times Y)^+$ a vector $\omega_{u, v} : A \rightarrow S$ can be defined in three ways.

Definition 4.1 (1n-semantics) For any $(x, y) \in X \times Y$ and $(u, v) \in (X \times Y)^+$ we set

$$\omega_{xu, yv} = D(\omega_{x, y}) \cdot \delta_x \cdot \omega_{u, v}. \quad (13)$$

In other words, for any $n \in \mathbb{N}$, $x_1, \dots, x_n \in X$ and $y_1, \dots, y_n \in Y$, we have that

$$\omega_{x_1 \dots x_n, y_1 \dots y_n} = D(\omega_{x_1, y_1}) \cdot \delta_{x_1} \cdot D(\omega_{x_2, y_2}) \cdot \delta_{x_2} \cdot \dots \cdot D(\omega_{x_{n-1}, y_{n-1}}) \cdot \delta_{x_{n-1}} \cdot \omega_{x_n, y_n}, \quad (14)$$

i.e., for every $a \in A$ the following is true

$$\omega_{x_1 \dots x_n, y_1 \dots y_n}(a) = \sum_{(a_1, \dots, a_{n-1}) \in A^{n-1}} \omega_{x_1, y_1}(a) \cdot \delta_{x_1}(a, a_1) \cdot \omega_{x_2, y_2}(a_1) \cdot \delta_{x_2}(a_1, a_2) \cdot \dots \cdot \omega_{x_{n-1}, y_{n-1}}(a_{n-2}) \cdot \delta_{x_{n-1}}(a_{n-2}, a_{n-1}) \cdot \omega_{x_n, y_n}(a_{n-1}). \quad (15)$$

The $1n$ -behavior of \mathcal{A} is the function $\llbracket \mathcal{A} \rrbracket_{1n} : (X \times Y)^* \rightarrow S$ defined by

$$\llbracket \mathcal{A} \rrbracket_{1n}(\varepsilon, \varepsilon) = \sigma \cdot \omega_{\varepsilon, \varepsilon} = \sum_{a \in A} \sigma(a) \quad (16)$$

and

$$\begin{aligned} \llbracket \mathcal{A} \rrbracket_{1n}(u, v) &= \sigma \cdot \omega_{u, v} = \sum_{a \in A} \sigma(a) \cdot \omega_{u, v}(a) \\ &= \sum_{(a, a_1, \dots, a_{n-1}) \in A^n} \sigma(a) \cdot \omega_{x_1, y_1}(a) \cdot \delta_{x_1}(a, a_1) \cdot \omega_{x_2, y_2}(a_1) \cdot \delta_{x_2}(a_1, a_2) \cdot \dots \cdot \omega_{x_{n-1}, y_{n-1}}(a_{n-2}) \cdot \delta_{x_{n-1}}(a_{n-2}, a_{n-1}) \cdot \omega_{x_n, y_n}(a_{n-1}), \end{aligned} \quad (17)$$

for each $(u, v) \in (X \times Y)^+$, $u = x_1 \dots x_n$, $v = y_1 \dots y_n$, for some $n \in \mathbb{N}$, $x_1, \dots, x_n \in X$ and $y_1, \dots, y_n \in Y$.

Definition 4.2 (n1-semantics) For any $(x, y) \in X \times Y$ and $(u, v) \in (X \times Y)^+$ we set

$$\omega_{ux, vy} = D(\omega_{u, v}) \cdot \delta_u \cdot \omega_{x, y}. \quad (18)$$

In other words, for each $n \in \mathbb{N}$, $x_1, \dots, x_n \in X$, and $y_1, \dots, y_n \in Y$ we have that

$$\omega_{x_1 \dots x_n, y_1 \dots y_n} = \omega_{x_1, y_1} \odot (\delta_{x_1} \cdot \omega_{x_2, y_2}) \odot (\delta_{x_1 x_2} \cdot \omega_{x_3, y_3}) \odot \dots \odot (\delta_{x_1 \dots x_{n-1}} \cdot \omega_{x_n, y_n}), \quad (19)$$

or equivalently, for every $a \in A$ we have

$$\begin{aligned} \omega_{x_1 \dots x_n, y_1 \dots y_n}(a) &= \sum_{(a_1, \dots, a_{n-1}) \in A^{n-1}} \omega_{x_1, y_1}(a) \cdot \delta_{x_1}(a, a_1) \cdot \omega_{x_2, y_2}(a_1) \cdot \delta_{x_1 x_2}(a, a_2) \cdot \omega_{x_3, y_3}(a_2) \cdot \dots \cdot \delta_{x_1 \dots x_{n-1}}(a, a_{n-1}) \cdot \omega_{x_n, y_n}(a_{n-1}). \end{aligned} \quad (20)$$

In this case, the $n1$ -behavior of \mathcal{A} is defined as the function $\llbracket \mathcal{A} \rrbracket_{n1} : (X \times Y)^* \rightarrow S$ given by

$$\llbracket \mathcal{A} \rrbracket_{n1}(\varepsilon, \varepsilon) = \sigma \cdot \omega_{\varepsilon, \varepsilon} = \sum_{a \in A} \sigma(a) \quad (21)$$

and

$$\begin{aligned} \llbracket \mathcal{A} \rrbracket_{n1}(u, v) &= \sigma \cdot \omega_{u, v} = \sum_{a \in A} \sigma(a) \cdot \omega_{u, v}(a) \\ &= \sum_{(a, a_1, \dots, a_{n-1}) \in A^n} \sigma(a) \cdot \omega_{x_1, y_1}(a) \cdot \delta_{x_1}(a, a_1) \cdot \omega_{x_2, y_2}(a_1) \cdot \delta_{x_1 x_2}(a, a_2) \cdot \omega_{x_3, y_3}(a_2) \cdot \dots \cdot \delta_{x_1 \dots x_{n-1}}(a, a_{n-1}) \cdot \omega_{x_n, y_n}(a_{n-1}), \end{aligned} \quad (22)$$

for each $(u, v) \in (X \times Y)^+$, $u = x_1 \dots x_n$, $v = y_1 \dots y_n$, for some $n \in \mathbb{N}$, $x_1, \dots, x_n \in X$ and $y_1, \dots, y_n \in Y$.

Definition 4.3 (Sequential semantics) The s -behavior of \mathcal{A} is the function $\llbracket \mathcal{A} \rrbracket_s : (X \times Y)^* \rightarrow S$ defined by

$$\llbracket \mathcal{A} \rrbracket_s(\varepsilon, \varepsilon) = \sigma \cdot \omega_{\varepsilon, \varepsilon} = \sum_{a \in A} \sigma(a) \quad (23)$$

and

$$\begin{aligned} \llbracket \mathcal{A} \rrbracket_s(u, v) &= \sum_{(a, a_1, \dots, a_n) \in A^{n+1}} \sigma(a) \cdot \omega_{x_1, y_1}(a) \cdot \delta_{x_1}(a, a_1) \cdot \omega_{x_2, y_2}(a_1) \cdot \delta_{x_2}(a_1, a_2) \cdot \dots \cdot \omega_{x_n, y_n}(a_{n-1}) \cdot \delta_{x_n}(a_{n-1}, a_n), \end{aligned} \quad (24)$$

for each $(u, v) \in (X \times Y)^+$, $u = x_1 \dots x_n$, $v = y_1 \dots y_n$, for some $n \in \mathbb{N}$, $x_1, \dots, x_n \in X$ and $y_1, \dots, y_n \in Y$.

If we define a function $\mu : A \times X \times Y \times A \rightarrow S$ by

$$\mu(a, x, y, b) = \omega(a, x, y) \cdot \delta(a, x, b), \quad (25)$$

for all $a, b \in A$, $x \in X$ and $y \in Y$, i.e., if we set

$$\mu_{x, y}(a, b) = \omega_{x, y}(a) \cdot \delta_x(a, b), \quad (26)$$

for all $a, b \in A$ and $(x, y) \in X \times Y$, we obtain a sequential weighted automaton $\mathcal{A}' = (A, X, Y, \sigma, \mu)$ such that

$\llbracket \mathcal{A}' \rrbracket = \llbracket \mathcal{A} \rrbracket_s$. For this reason this semantics is called sequential.

Let us note that $\mu_{x,y} = D(\omega_{x,y}) \cdot \delta_x$, for all $x \in X$ and $y \in Y$, and therefore,

$$\begin{aligned} \mu_{x_1 \dots x_n, y_1 \dots y_n} &= \\ &= D(\omega_{x_1, y_1}) \cdot \delta_{x_1} \cdot D(\omega_{x_2, y_2}) \cdot \delta_{x_2} \cdot \dots \cdot D(\omega_{x_n, y_n}) \cdot \delta_{x_n}, \end{aligned} \quad (27)$$

for any $n \in \mathbb{N}$ and $x_1, \dots, x_n \in X$, $y_1, \dots, y_n \in Y$.

Example 4.1 Let $S = ([0, 1], \vee, \wedge, 0, 1)$ be the Gödel semiring, and $\mathcal{A} = (A, X, Y, \sigma, \delta, \omega)$ a Mealy-type weighted automaton over S with $|A| = 2$, $X = \{0\}$, $Y = \{0, 1\}$, and

$$\begin{aligned} \sigma &= \begin{bmatrix} 1 & 0 \end{bmatrix}, \quad \delta_0 = \begin{bmatrix} 0.7 & 0.5 \\ 0 & 0.8 \end{bmatrix}, \\ \omega_{0,0} &= \begin{bmatrix} 0.6 & 0.4 \end{bmatrix}, \quad \omega_{0,1} = \begin{bmatrix} 0.2 & 0.7 \end{bmatrix}. \end{aligned}$$

It is easy to check that

$$\begin{aligned} \llbracket \mathcal{A} \rrbracket_{1n}(000, 010) &= \llbracket \mathcal{A} \rrbracket_s(000, 010) = 0.4 \\ &\neq 0.5 = \llbracket \mathcal{A} \rrbracket_{n1}(000, 010). \end{aligned}$$

Therefore, both the $1n$ -semantics and the sequential semantics differ from the $n1$ -semantics.

Example 4.2 Again, let S be the Gödel semiring, and let $\mathcal{A} = (A, X, Y, \sigma, \delta, \omega)$ be a Mealy-type weighted automaton over S with $|A| = 2$, $X = \{0, 1\}$, $Y = \{0\}$, and

$$\begin{aligned} \sigma &= \begin{bmatrix} 1 & 0 \end{bmatrix}, \quad \delta_0 = \begin{bmatrix} 0.7 & 0.5 \\ 0 & 0.8 \end{bmatrix}, \quad \delta_1 = \begin{bmatrix} 0.3 & 1 \\ 0.2 & 0 \end{bmatrix}, \\ \omega_{0,0} &= \begin{bmatrix} 0.6 & 0.4 \end{bmatrix}, \quad \omega_{1,0} = \begin{bmatrix} 0.2 & 0.7 \end{bmatrix}. \end{aligned}$$

Then

$$\llbracket \mathcal{A} \rrbracket_{1n}(01, 00) = 0.5 \neq 0.2 = \llbracket \mathcal{A} \rrbracket_s(01, 00),$$

and hence, the $1n$ -semantics and the sequential semantics may also be different.

5 Moore-type weighted automata

A Moore-type weighted automaton over a semiring S is a tuple $\mathcal{A} = (A, X, Y, \sigma^A, \delta^A, \omega^A)$, where everything is the same as in the definition of a Mealy-type weighted automaton except the weighted output function, for which we assume that $\omega^A : A \times Y \rightarrow S$.

Here, we define $\omega_{\varepsilon, \varepsilon} : A \rightarrow S$ by $\omega_{\varepsilon, \varepsilon}(a) = 1$, for each $a \in A$, and for any $(x, y) \in X \times Y$ we define $\omega_{x,y} : A \rightarrow L$ by $\omega_{x,y} = \delta_x \cdot \omega_y$. For $(u, v) \in (X \times Y)^+$ we can define a vector $\omega_{u,v} : A \rightarrow S$ in two ways.

Definition 5.1 (1n-semantics) For each $(x, y) \in X \times Y$ and $(u, v) \in (X \times Y)^+$ we set

$$\omega_{xu, yv} = \delta_x \cdot D(\omega_y) \cdot \omega_{u,v}. \quad (28)$$

In other words, for each $n \in \mathbb{N}$, $x_1, \dots, x_n \in X$ and $y_1, \dots, y_n \in Y$, we have that

$$\omega_{x_1 \dots x_n, y_1 \dots y_n} = \delta_{x_1} \cdot D(\omega_{y_1}) \cdot \delta_{x_2} \cdot D(\omega_{y_2}) \cdot \dots \cdot \delta_{x_n} \cdot \omega_{y_n}, \quad (29)$$

i.e., for every $a \in A$ we have

$$\begin{aligned} \omega_{x_1 \dots x_n, y_1 \dots y_n}(a) &= \sum_{(a_1, \dots, a_n) \in A^n} \delta_{x_1}(a, a_1) \cdot \omega_{y_1}(a_1) \cdot \\ &\cdot \delta_{x_2}(a_1, a_2) \cdot \omega_{y_2}(a_2) \cdot \dots \cdot \delta_{x_n}(a_{n-1}, a_n) \cdot \omega_{y_n}(a_n). \end{aligned} \quad (30)$$

The $1n$ -behavior of \mathcal{A} is the function $\llbracket \mathcal{A} \rrbracket_{1n} : (X \times Y)^* \rightarrow L$ defined by

$$\llbracket \mathcal{A} \rrbracket_{1n}(\varepsilon, \varepsilon) = \sigma \cdot \omega_{\varepsilon, \varepsilon} = \sum_{a \in A} \sigma(a) \quad (31)$$

and

$$\begin{aligned} \llbracket \mathcal{A} \rrbracket_{1n}(u, v) &= \sigma \cdot \omega_{u,v} = \sum_{a \in A} \sigma(a) \cdot \omega_{u,v}(a) \\ &= \sum_{(a, a_1, \dots, a_n) \in A^{n+1}} \sigma(a) \cdot \delta_{x_1}(a, a_1) \cdot \omega_{y_1}(a_1) \cdot \\ &\cdot \delta_{x_2}(a_1, a_2) \cdot \omega_{y_2}(a_2) \cdot \dots \cdot \delta_{x_n}(a_{n-1}, a_n) \cdot \omega_{y_n}(a_n), \end{aligned} \quad (32)$$

for each $(u, v) \in (X \times Y)^+$, $u = x_1 \dots x_n$, $v = y_1 \dots y_n$, for some $n \in \mathbb{N}$, $x_1, \dots, x_n \in X$ and $y_1, \dots, y_n \in Y$.

Note that a slightly different definition of $1n$ -semantics for Moore-type fuzzy finite automata was given by Li and Pedrycz in [3].

Definition 5.2 (n1-semantics) For each $(x, y) \in X \times Y$ and $(u, v) \in (X \times Y)^+$ we set

$$\omega_{ux, vy} = D(\omega_{u,v}) \cdot \delta_{ux} \cdot \omega_y. \quad (33)$$

In other words, for each $n \in \mathbb{N}$, $x_1, \dots, x_n \in X$, and $y_1, \dots, y_n \in Y$ we have that

$$\begin{aligned} \omega_{x_1 \dots x_n, y_1 \dots y_n} &= \\ &= (\delta_{x_1} \cdot \omega_{y_1}) \odot (\delta_{x_1 x_2} \cdot \omega_{y_2}) \odot \dots \odot (\delta_{x_1 \dots x_n} \cdot \omega_{y_n}), \end{aligned} \quad (34)$$

which means that

$$\begin{aligned} \omega_{x_1 \dots x_n, y_1 \dots y_n}(a) &= \sum_{(a_1, \dots, a_n) \in A^n} \delta_{x_1}(a, a_1) \cdot \omega_{y_1}(a_1) \cdot \\ &\cdot \delta_{x_1 x_2}(a, a_2) \cdot \omega_{y_2}(a_2) \cdot \dots \cdot \delta_{x_1 \dots x_n}(a, a_n) \cdot \omega_{y_n}(a_n), \end{aligned} \quad (35)$$

for every $a \in A$. Now, the $n1$ -behavior of \mathcal{A} is defined as the function $\llbracket \mathcal{A} \rrbracket_{n1} : (X \times Y)^* \rightarrow S$ given by

$$\llbracket \mathcal{A} \rrbracket_{n1}(\varepsilon, \varepsilon) = \sigma \cdot \omega_{\varepsilon, \varepsilon} = \sum_{a \in A} \sigma(a) \quad (36)$$

and

$$\begin{aligned} \llbracket \mathcal{A} \rrbracket_{n1}(u, v) &= \sigma \cdot \omega_{u,v} = \sum_{a \in A} \sigma(a) \cdot \omega_{u,v}(a) \\ &= \sum_{(a, a_1, \dots, a_n) \in A^{n+1}} \sigma(a) \cdot \delta_{x_1}(a, a_1) \cdot \omega_{y_1}(a_1) \cdot \\ &\quad \cdot \delta_{x_1 x_2}(a, a_2) \cdot \omega_{y_2}(a_2) \cdot \dots \cdot \delta_{x_1 \dots x_n}(a, a_n) \cdot \omega_{y_n}(a_n), \end{aligned} \quad (37)$$

for every $(u, v) \in (X \times Y)^+$, $u = x_1 \dots x_n$, $v = y_1 \dots y_n$, for some $n \in \mathbb{N}$, $x_1, \dots, x_n \in X$ and $y_1, \dots, y_n \in Y$.

6 Equivalence of sequential, Moore-type and Mealy-type weighted automata

Two weighted finite automata with output of any type (sequential, Mealy-type or Moore type) are *equivalent* if they have equal behaviors (with respect to the considered semantics). In this section we prove theorems on the equivalence of sequential, Mealy-type and Moore-type weighted finite automata with respect to various semantics.

First we prove that every Mealy-type weighted automaton \mathcal{A} can be converted into a sequential weighted automaton which is equivalent to \mathcal{A} with respect to sequential semantics on \mathcal{A} .

Theorem 6.1 *For any Mealy-type weighted automaton $\mathcal{A} = (A, X, Y, \sigma^A, \delta^A, \omega^A)$ there exists a sequential weighted automaton $\mathcal{B} = (B, X, Y, \sigma^B, \delta^B, \omega^B)$ such that*

$$\llbracket \mathcal{A} \rrbracket_s = \llbracket \mathcal{B} \rrbracket.$$

In addition, \mathcal{B} can be chosen so that $|\mathcal{B}| \leq |\mathcal{A}|$.

Proof Set $B = A$ and define $\mu^B : B \times X \times Y \times B \rightarrow S$ and $\sigma^B : B \rightarrow S$ by $\sigma^B = \sigma^A$ and

$$\mu(a, x, y, b) = \omega(a, x, y) \cdot \delta(a, x, b),$$

for all $a, b \in A$, $x \in X$ and $y \in Y$. Then it is easy to check that $\llbracket \mathcal{A} \rrbracket_s = \llbracket \mathcal{B} \rrbracket$. \square

Next, we show that every Moore-type weighted automaton \mathcal{A} can be converted into a sequential weighted automaton which is equivalent to \mathcal{A} with respect to $1n$ -semantics on \mathcal{A} .

Theorem 6.2 *For any Moore-type weighted automaton $\mathcal{A} = (A, X, Y, \sigma^A, \delta^A, \omega^A)$ there exists a sequential weighted automaton $\mathcal{B} = (B, X, Y, \sigma^B, \delta^B, \omega^B)$ such that*

$$\llbracket \mathcal{A} \rrbracket_{1n} = \llbracket \mathcal{B} \rrbracket.$$

In addition, \mathcal{B} can be chosen so that $|\mathcal{B}| \leq |\mathcal{A}|$.

Proof Set $B = A$ and define $\mu^B : B \times X \times Y \times B \rightarrow S$ and $\sigma^B : B \rightarrow S$ by $\sigma^B = \sigma^A$ and

$$\mu(a, x, y, b) = \delta(a, x, b) \cdot \omega(b, y),$$

for all $a, b \in A$, $x \in X$ and $y \in Y$. Then $\llbracket \mathcal{A} \rrbracket_{1n} = \llbracket \mathcal{B} \rrbracket$. \square

On the other hand, the next theorem shows that any sequential weighted automaton \mathcal{A} can be converted to a Moore-type weighted automaton \mathcal{B} which is equivalent to \mathcal{A} with respect to $1n$ -semantics on \mathcal{B} .

Theorem 6.3 *For any sequential weighted automaton $\mathcal{A} = (A, X, Y, \sigma^A, \mu^A)$ there exists a Moore-type weighted automaton $\mathcal{B} = (B, X, Y, \sigma^B, \delta^B, \omega^B)$ such that*

$$\llbracket \mathcal{A} \rrbracket = \llbracket \mathcal{B} \rrbracket_{1n}.$$

In addition, \mathcal{B} can be chosen so that $|\mathcal{B}| \leq |\mathcal{A}| \cdot |Y|$.

Proof Set $B = A \times Y$ and fix an arbitrary $y_0 \in Y$. Define $\sigma^B : B \rightarrow S$, $\delta^B : B \times X \times B \rightarrow S$ and $\omega^B : B \times Y \rightarrow S$ as follows: For $b, b_1, b_2 \in B$, $x \in X$ and $y \in Y$ we set

$$\sigma^B(b) = \begin{cases} \sigma^A(a) & \text{if } b = (a, y_0), \text{ for some } a \in A, \\ 0 & \text{otherwise,} \end{cases}$$

$$\delta^B(b_1, x, b_2) = \mu^A(a_1, x, y_2, a_2), \quad \text{if } b_1 = (a_1, y_1), \\ b_2 = (a_2, y_2), \text{ for some } a_1, a_2 \in A, y_1, y_2 \in Y,$$

$$\omega^B(b, y) = \begin{cases} 1 & \text{if } b = (a, y), \text{ for some } a \in A, \\ 0 & \text{otherwise.} \end{cases}$$

Then $\mathcal{B} = (B, X, Y, \sigma^B, \delta^B, \omega^B)$ is a Moore-type weighted automaton. We are going to prove that \mathcal{A} is equivalent to \mathcal{B} with respect to the $1n$ -semantics of \mathcal{B} .

Take an arbitrary $(u, v) \in (X \times Y)^+$, where $u = x_1 \dots x_n$, $v = y_1 \dots y_n$, for $n \in \mathbb{N}$, $x_1, \dots, x_n \in X$, $y_1, \dots, y_n \in Y$. Consider any $(b_0, b_1, \dots, b_n) \in B^{n+1}$ and the product

$$\begin{aligned} &\sigma^B(b_0) \cdot \delta_{x_1}^B(b_0, b_1) \cdot \omega_{y_1}^B(b_1) \cdot \delta_{x_2}^B(b_1, b_2) \cdot \omega_{y_2}^B(b_2) \cdot \\ &\quad \dots \cdot \delta_{x_n}^B(b_{n-1}, b_n) \cdot \omega_{y_n}^B(b_n). \end{aligned} \quad (38)$$

If for each $i \in \{0, 1, \dots, n\}$ we have that

$$b_i = (a_i, y_i), \quad \text{for some } a_i \in A, \quad (39)$$

then

$$\sigma^B(b_0) = \sigma^A(a_0), \quad \omega_{y_i}^B(b_i) = 1,$$

$$\delta_{x_i}^B(b_{i-1}, b_i) = \mu_{x_i, y_i}^A(a_{i-1}, a_i),$$

for each $i \in \{1, \dots, n\}$, and the product (38) becomes

$$\sigma^A(a_0) \cdot \mu_{x_1, y_1}^A(a_0, a_1) \cdot \mu_{x_2, y_2}^A(a_1, a_2) \cdot \dots \cdot \mu_{x_n, y_n}^A(a_{n-1}, a_n).$$

Otherwise, if there exists $i \in \{0, 1, \dots, n\}$ such that b_i can not be written in the form (39), then we have that

$\omega_{y_i}^B(b_i) = 0$ (for $i \geq 1$) or $\sigma^B(b_0) = 0$ (for $i = 0$), and the whole product (38) is equal to 0.

Note also that there is a one-to-one correspondence between all $(n+1)$ -tuples $(a_0, a_1, \dots, a_n) \in A^{n+1}$ and all $(n+1)$ -tuples $(b_0, b_1, \dots, b_n) \in B^{n+1}$ satisfying (39), which implies

$$\begin{aligned} \llbracket \mathcal{B} \rrbracket_{1n}(u, v) &= \sum_{(b_0, b_1, \dots, b_n) \in B^{n+1}} \sigma^B(b_0) \cdot \delta_{x_1}^B(b_0, b_1) \cdot \omega_{y_1}^B(b_1) \cdot \\ &\quad \cdot \delta_{x_2}^B(b_1, b_2) \cdot \omega_{y_2}^B(b_2) \cdot \dots \cdot \delta_{x_n}^B(b_{n-1}, b_n) \cdot \omega_{y_n}^B(b_n) \\ &= \sum_{(a_0, a_1, \dots, a_n) \in A^{n+1}} \sigma^A(a_0) \cdot \mu_{x_1, y_1}^A(a_0, a_1) \cdot \mu_{x_2, y_2}^A(a_1, a_2) \cdot \\ &\quad \cdot \dots \cdot \mu_{x_n, y_n}^A(a_{n-1}, a_n) = \llbracket \mathcal{A} \rrbracket(u, v), \end{aligned}$$

and hence, $\llbracket \mathcal{B} \rrbracket_{1n} = \llbracket \mathcal{A} \rrbracket$. Clearly, $|\mathcal{B}| \leq |\mathcal{A}| \cdot |Y|$. \square

Then we show that any Mealy-type weighted automaton \mathcal{A} can be converted into a Moore-type weighted automaton \mathcal{B} such that \mathcal{A} and \mathcal{B} are equivalent both with respect to $1n$ -semantics and $n1$ -semantics.

Theorem 6.4 *For every Mealy-type weighted automaton $\mathcal{A} = (A, X, Y, \sigma^A, \delta^A, \omega^A)$ there exists a Moore-type weighted automaton $\mathcal{B} = (B, X, Y, \sigma^B, \delta^B, \omega^B)$ such that*

$$\llbracket \mathcal{A} \rrbracket_{1n} = \llbracket \mathcal{B} \rrbracket_{1n} \quad \text{and} \quad \llbracket \mathcal{A} \rrbracket_{n1} = \llbracket \mathcal{B} \rrbracket_{n1}.$$

In addition, \mathcal{B} can be chosen so that $|\mathcal{B}| \leq |\mathcal{A}| \cdot (|X| + 1)$.

Proof Set $B = A \cup A \times X$. Let us define $\sigma^B : B \rightarrow S$, $\delta^B : B \times X \times B \rightarrow S$ and $\omega^B : B \times Y \rightarrow S$ as follows: For $b, b_1, b_2 \in B$, $x \in X$ and $y \in Y$ we set

$$\begin{aligned} \sigma^B(b) &= \begin{cases} \sigma^A(a) & \text{if } b = a \in A, \\ 0 & \text{otherwise,} \end{cases} \\ \delta^B(b_1, x, b_2) &= \begin{cases} 1 & \text{if } b_1 = a \in A, \\ & b_2 = (a, x) \in A \times X, \\ \delta^A(a_1, x_1, a_2) & \text{if } b_1 = (a_1, x_1) \in A \times X, \\ & b_2 = (a_2, x) \in A \times X, \\ 0 & \text{otherwise,} \end{cases} \\ \omega^B(b, y) &= \begin{cases} \omega^A(a, x, y) & \text{if } b = (a, x) \in A \times X, \\ 0 & \text{otherwise.} \end{cases} \end{aligned}$$

Then $\mathcal{B} = (B, X, Y, \sigma^B, \delta^B, \omega^B)$ is a Moore-type weighted automaton.

Take an arbitrary $(u, v) \in (X \times Y)^+$, where $u = x_1 \dots x_n$, $v = y_1 \dots y_n$, for $n \in \mathbb{N}$, $x_1, \dots, x_n \in X$, $y_1, \dots, y_n \in Y$. Consider any $(b_0, b_1, \dots, b_n) \in B^{n+1}$ and the product

$$\begin{aligned} &\sigma^B(b_0) \cdot \delta_{x_1}^B(b_0, b_1) \cdot \omega_{y_1}^B(b_1) \cdot \delta_{x_2}^B(b_1, b_2) \cdot \omega_{y_2}^B(b_2) \cdot \\ &\quad \cdot \dots \cdot \delta_{x_n}^B(b_{n-1}, b_n) \cdot \omega_{y_n}^B(b_n). \end{aligned} \quad (40)$$

Suppose that

$$b_0 = a_0, \quad \text{for some } a_0 \in A, \quad (41)$$

and for any $i \in \{1, \dots, n\}$ suppose that

$$b_i = (a_{i-1}, x_i), \quad \text{for some } a_{i-1} \in A. \quad (42)$$

Then

$$\sigma^B(b_0) = \sigma^A(a_0), \quad (43)$$

$$\omega_{y_i}^B(b_i) = \omega_{x_i, y_i}^A(a_{i-1}), \quad \text{for } i \in \{1, \dots, n\}, \quad (44)$$

$$\delta_{x_1}^B(b_0, b_1) = 1, \quad (45)$$

$$\delta_{x_i}^B(b_{i-1}, b_i) = \delta_{x_{i-1}}^A(a_{i-2}, a_{i-1}), \quad \text{for } i \in \{2, \dots, n\}, \quad (46)$$

and the product (40) becomes

$$\begin{aligned} &\sigma^A(a_0) \cdot \omega_{x_1, y_1}^A(a_0) \cdot \delta_{x_1}^A(a_0, a_1) \cdot \omega_{x_2, y_2}^A(a_1) \cdot \delta_{x_2}^A(a_1, a_2) \cdot \\ &\quad \cdot \dots \cdot \delta_{x_{n-1}}^A(a_{n-2}, a_{n-1}) \cdot \omega_{x_n, y_n}^A(a_{n-1}). \end{aligned}$$

On the other hand, if $b_0 \in A \times X$ or if b_i can not be written in the form (42), for some $i \in \{1, \dots, n\}$, i.e., if $b_i \in A$ or $b_i = (a, x) \in A \times X$ such that $x \neq x_i$, then $\sigma^B(b_0) = 0$ or $\delta_{x_i}^B(b_{i-1}, b_i) = 0$, and in both cases the whole product (40) is equal to 0.

Since to each $(n+1)$ -tuple $(a_0, \dots, a_n) \in A^{n+1}$ corresponds exactly one $(n+1)$ -tuple $(b_0, \dots, b_n) \in B^{n+1}$ satisfying (41) and (42), we have that

$$\begin{aligned} \llbracket \mathcal{B} \rrbracket_{1n}(u, v) &= \sum_{(b_0, \dots, b_n) \in B^{n+1}} \sigma^B(b_0) \cdot \delta_{x_1}^B(b_0, b_1) \cdot \omega_{y_1}^B(b_1) \cdot \\ &\quad \cdot \delta_{x_2}^B(b_1, b_2) \cdot \omega_{y_2}^B(b_2) \cdot \dots \cdot \delta_{x_n}^B(b_{n-1}, b_n) \cdot \omega_{y_n}^B(b_n) \\ &= \sum_{(a_0, \dots, a_n) \in A^{n+1}} \sigma^A(a_0) \cdot \omega_{x_1, y_1}^A(a_0) \cdot \delta_{x_1}^A(a_0, a_1) \cdot \omega_{x_2, y_2}^A(a_1) \cdot \\ &\quad \cdot \delta_{x_2}^A(a_1, a_2) \cdot \dots \cdot \delta_{x_{n-1}}^A(a_{n-2}, a_{n-1}) \cdot \omega_{x_n, y_n}^A(a_{n-1}) \\ &= \llbracket \mathcal{A} \rrbracket_{1n}(u, v), \end{aligned}$$

and hence, $\llbracket \mathcal{B} \rrbracket_{1n} = \llbracket \mathcal{A} \rrbracket_{1n}$.

Next, we prove that $\llbracket \mathcal{B} \rrbracket_{n1} = \llbracket \mathcal{A} \rrbracket_{n1}$. Consider again an arbitrary $(b_0, b_1, \dots, b_n) \in B^{n+1}$ and the product

$$\begin{aligned} &\sigma^B(b_0) \cdot \delta_{x_1}^B(b_0, b_1) \cdot \omega_{y_1}^B(b_1) \cdot \delta_{x_1 x_2}^B(b_0, b_2) \cdot \omega_{y_2}^B(b_2) \cdot \\ &\quad \cdot \dots \cdot \delta_{x_1 \dots x_n}^B(b_0, b_n) \cdot \omega_{y_n}^B(b_n). \end{aligned} \quad (47)$$

Suppose again that (41) and (42) hold. Then we have that (43), (44) and (45) also hold. Now, take an arbitrary $j \in \{2, \dots, n\}$ and $(b'_1, \dots, b'_{j-1}) \in B^{j-1}$, and consider the product

$$\delta_{x_1}^B(b_0, b'_1) \cdot \delta_{x_2}^B(b'_1, b'_2) \cdot \dots \cdot \delta_{x_j}^B(b'_{j-1}, b_j). \quad (48)$$

If

$$b'_1 = (a_0, x_1), \quad (49)$$

and if for any $k \in \{2, \dots, j-1\}$ we have that

$$b'_k = (a'_{k-1}, x_k), \text{ for some } a'_{k-1} \in A, \quad (50)$$

then

$$\begin{aligned} \delta_{x_1}^B(b_0, b'_1) &= 1, \\ \delta_{x_2}^B(b'_1, b'_2) &= \delta_{x_1}^A(a_0, a'_1), \\ \delta_{x_k}^B(b'_{k-1}, b'_k) &= \delta_{x_{k-1}}^A(a'_{k-2}, a'_{k-1}), \text{ for } k \in \{3, \dots, j-1\}, \\ \delta_{x_j}^B(b'_{j-1}, b_j) &= \delta_{x_{j-1}}^A(a'_{j-2}, a_{j-1}), \end{aligned}$$

and consequently, the product (48) becomes

$$\delta_{x_1}^A(a_0, a'_1) \cdot \delta_{x_2}^A(a'_1, a'_2) \cdot \dots \cdot \delta_{x_{j-1}}^A(a'_{j-2}, a_{j-1}).$$

Otherwise, if $b'_1 \neq (a_0, x_1)$ or there is $k \in \{2, \dots, j-1\}$ such that (42) does not hold, then $\delta_{x_k}^B(b'_{k-1}, b'_k) = 0$ and the product (48) is also equal to 0. Therefore

$$\begin{aligned} \delta_{x_1 \dots x_j}^B(b_0, b_j) &= \\ &= \sum_{(b'_1, \dots, b'_{j-1}) \in B^{j-1}} \delta_{x_1}^B(b_0, b'_1) \cdot \delta_{x_2}^B(b'_1, b'_2) \cdot \dots \cdot \delta_{x_j}^B(b'_{j-1}, b_j) \\ &= \sum_{(a'_1, \dots, a'_{j-2}) \in A^{j-2}} \delta_{x_1}^A(a_0, a'_1) \cdot \delta_{x_2}^A(a'_1, a'_2) \cdot \dots \cdot \delta_{x_{j-1}}^A(a'_{j-2}, a_{j-1}) \\ &= \delta_{x_1 \dots x_{j-1}}^A(a_0, a_{j-1}), \end{aligned}$$

for any $j \in \{2, \dots, n\}$, which means that the product (47) becomes

$$\begin{aligned} \sigma^A(a_0) \cdot \omega_{x_1, y_1}^A(a_0) \cdot \delta_{x_1}^A(a_0, a_1) \cdot \omega_{x_2, y_2}^A(a_1) \cdot \delta_{x_1 x_2}^A(a_0, a_2) \cdot \\ \cdot \dots \cdot \delta_{x_1 \dots x_{n-1}}^A(a_0, a_{n-1}) \cdot \omega_{x_n, y_n}^A(a_{n-1}). \end{aligned}$$

Next, if $b_0 \in A \times X$ or there exists $i \in \{1, \dots, n\}$ such that $b_i = (a, x) \in A \times X$ with $x \neq x_i$, then $\delta_{x_i}^B(b, b_i) = 0$, for any $b \in B$, whence $\delta_{x_1 \dots x_i}^B(b_0, b_i) = 0$, which implies that the product (47) is equal to 0. Now, we conclude that

$$\begin{aligned} \llbracket \mathcal{B} \rrbracket_{1n}(u, v) &= \sum_{(b_0, \dots, b_n) \in B^{n+1}} \sigma^B(b_0) \cdot \delta_{x_1}^B(b_0, b_1) \cdot \omega_{y_1}^B(b_1) \cdot \\ &\cdot \delta_{x_1 x_2}^B(b_0, b_2) \cdot \omega_{y_2}^B(b_2) \cdot \dots \cdot \delta_{x_1 \dots x_n}^B(b_0, b_n) \cdot \omega_{y_n}^B(b_n) \\ &= \sum_{(a_0, \dots, a_{n-1}) \in A^n} \sigma^A(a_0) \cdot \omega_{x_1, y_1}^A(a_0) \cdot \delta_{x_1}^A(a_0, a_1) \cdot \\ &\cdot \omega_{x_2, y_2}^A(a_1) \cdot \delta_{x_1 x_2}^A(a_0, a_2) \cdot \omega_{x_3, y_3}^A(a_2) \cdot \\ &\cdot \dots \cdot \delta_{x_1 \dots x_{n-1}}^A(a_0, a_{n-1}) \cdot \omega_{x_n, y_n}^A(a_{n-1}) \\ &= \llbracket \mathcal{A} \rrbracket_{1n}(u, v). \end{aligned}$$

Therefore, $\llbracket \mathcal{B} \rrbracket_{1n} = \llbracket \mathcal{A} \rrbracket_{1n}$. \square

We also prove that any Moore-type weighted automaton \mathcal{A} can be converted into a Mealy-type weighted automaton \mathcal{B} such that \mathcal{A} and \mathcal{B} are equivalent with respect to $1n$ -semantics.

Theorem 6.5 *For every Moore-type weighted automaton $\mathcal{A} = (A, X, Y, \sigma^A, \delta^A, \omega^A)$ there is a Mealy-type weighted automaton $\mathcal{B} = (B, X, Y, \sigma^B, \delta^B, \omega^B)$ such that*

$$\llbracket \mathcal{A} \rrbracket_{1n} = \llbracket \mathcal{B} \rrbracket_{1n}.$$

In addition, \mathcal{B} can be chosen so that $|\mathcal{B}| \leq |\mathcal{A}|^2$.

Proof Let $B = A \times A$ and let $\sigma^B : B \rightarrow S$, $\delta^B : B \times X \times B \rightarrow S$ and $\omega^B : B \times X \times Y \rightarrow S$ be defined as follows: For $b, b_1, b_2 \in B$, $x \in X$ and $y \in Y$ we set

$$\sigma^B(b) = \sigma^A(a), \quad \text{if } b = (a, a'), \text{ for some } a, a' \in A,$$

$$\delta^B(b_1, x, b_2) = \begin{cases} 1 & \text{if } b_1 = (a_1, a_2), b_2 = (a_2, a_3), \\ & \text{for some } a_1, a_2, a_3 \in A, \\ 0 & \text{otherwise,} \end{cases}$$

$$\omega^B(b, x, y) = \delta_x^A(a_1, a_2) \cdot \omega_y(a_2), \quad \text{if } b = (a_1, a_2),$$

for some $a_1, a_2 \in A$,

Then $\mathcal{B} = (B, X, Y, \sigma^B, \delta^B, \omega^B)$ is a Mealy-type weighted automaton. To prove that $\llbracket \mathcal{A} \rrbracket_{1n} = \llbracket \mathcal{B} \rrbracket_{1n}$ take an arbitrary $(u, v) \in (X \times Y)^+$, where $u = x_1 \dots x_n$, $v = y_1 \dots y_n$, for some $n \in \mathbb{N}$, $x_1, \dots, x_n \in X$, $y_1, \dots, y_n \in Y$. Consider any $(b_0, b_1, \dots, b_{n-1}) \in B^n$ and the product

$$\begin{aligned} \sigma^B(b_0) \cdot \omega_{x_1, y_1}^B(b_0) \cdot \delta_{x_1}^B(b_0, b_1) \cdot \omega_{x_2, y_2}^B(b_1) \cdot \\ \cdot \dots \cdot \delta_{x_n}^B(b_{n-1}, b_n) \cdot \omega_{x_n, y_n}^B(b_{n-1}). \end{aligned} \quad (51)$$

Suppose that there is $(a_0, a_1, \dots, a_n) \in A^{n+1}$ such that

$$b_{i-1} = (a_{i-1}, a_i), \quad \text{for each } i \in \{1, \dots, n\}. \quad (52)$$

Then we have that

$$\begin{aligned} \sigma^B(b_0) &= \sigma^A(a_0), \\ \delta_{x_i}^B(b_{i-1}, b_i) &= 1, \quad \text{for each } i \in \{1, \dots, n-1\}, \\ \omega_{x_i, y_i}^B(b_{i-1}) &= \delta_{x_i}^A(a_{i-1}, a_i) \cdot \omega_{y_i}^A(a_i), \end{aligned}$$

and consequently, the product (51) becomes

$$\begin{aligned} \sigma^A(a_0) \cdot \delta_{x_1}^A(a_0, a_1) \cdot \omega_{y_1}^A(a_1) \cdot \delta_{x_2}^A(a_1, a_2) \cdot \\ \cdot \omega_{y_2}^A(a_2) \cdot \dots \cdot \delta_{x_n}^A(a_{n-1}, a_n) \cdot \omega_{y_n}^A(a_n). \end{aligned}$$

On the other hand, if there is $i \in \{1, \dots, n-1\}$ such that $b_{i-1} = (a'_1, a'_2)$, $b_i = (a''_1, a''_2)$ and $a'_2 \neq a''_1$, then we obtain that $\delta_{x_i}^B(b_{i-1}, b_i) = 0$, and the product (51) is equal to 0.

Since for any $(n+1)$ -tuple $(a_0, a_1, \dots, a_n) \in A^{n+1}$ there exists a unique n -tuple $(b_0, b_1, \dots, b_{n-1}) \in B^n$ such that (52) holds, we have that

$$\begin{aligned} \llbracket \mathcal{B} \rrbracket_{1n}(u, v) &= \sum_{(b_0, b_1, \dots, b_{n-1}) \in B^n} \sigma^B(b_0) \cdot \omega_{x_1, y_1}^B(b_0) \cdot \delta_{x_1}^B(b_0, b_1) \cdot \\ &\cdot \omega_{x_2, y_2}^B(b_1) \cdot \dots \cdot \delta_{x_n}^B(b_{n-1}, b_n) \cdot \omega_{x_n, y_n}^B(b_{n-1}) \\ &= \sum_{(a_0, a_1, \dots, a_n) \in A^{n+1}} \sigma^A(a_0) \cdot \delta_{x_1}^A(a_0, a_1) \cdot \omega_{y_1}^A(a_1) \cdot \\ &\cdot \delta_{x_2}^A(a_1, a_2) \cdot \omega_{y_2}^A(a_2) \cdot \dots \cdot \delta_{x_n}^A(a_{n-1}, a_n) \cdot \omega_{y_n}^A(a_n) \\ &= \llbracket \mathcal{A} \rrbracket_{1n}(u, v), \end{aligned}$$

and hence, $\llbracket \mathcal{B} \rrbracket_{1n} = \llbracket \mathcal{A} \rrbracket_{1n}$. \square

Finally, we show that under certain conditions a sequential weighted finite automaton \mathcal{A} can be converted to a Mealy-type weighted finite automaton \mathcal{B} equivalent to \mathcal{A} with respect to the sequential semantics on \mathcal{B} .

Theorem 6.6 *Given a sequential weighted automaton $\mathcal{A} = (A, X, Y, \sigma^A, \mu^A)$. If there exists $p \in \mathbb{N}$ such that $(pk)s = s$, for any $s \in \text{Im}(\mu^A)$, where $k = |X| \cdot |Y|$, then there is a Mealy-type weighted automaton $\mathcal{B} = (B, X, Y, \sigma^B, \delta^B, \omega^B)$ such that*

$$\llbracket \mathcal{A} \rrbracket = \llbracket \mathcal{B} \rrbracket_s.$$

In addition, \mathcal{B} can be chosen so that $|\mathcal{B}| \leq |\mathcal{A}| \cdot |X| \cdot |Y|$.

Proof Set $B = A \times X \times Y$. Let us define $\sigma^B : B \rightarrow S$, $\delta^B : B \times X \times B \rightarrow S$ and $\omega^B : B \times X \times Y \rightarrow S$ as follows: For $b, b_1, b_2 \in B$, $x \in X$ and $y \in Y$ we set

$$\begin{aligned} \sigma^B(b) &= p \sigma^A(a), & \text{if } b = (a, x_1, y_1), \text{ for some } a \in A, \\ & & x_1 \in X \text{ and } y_1 \in Y, \\ \delta^B(b_1, x, b_2) &= \mu^A(a_1, x, y_1, a_2), & \text{if } b_1 = (a_1, x_1, y_1), \\ & & b_2 = (a_2, x_2, y_2), \text{ for some } a_1, a_2 \in A, \\ & & x_1, x_2 \in X, y_1, y_2 \in Y, \end{aligned}$$

$$\omega^B(b, x, y) = \begin{cases} 1 & \text{if } b = (a, x, y), \text{ for some } a \in A, \\ 0 & \text{otherwise.} \end{cases}$$

Then $\mathcal{B} = (B, X, Y, \sigma^B, \delta^B, \omega^B)$ is a Mealy-type weighted automaton. We will show that \mathcal{A} is equivalent to \mathcal{B} with respect to the sequential semantics of \mathcal{B} .

Take an arbitrary $(u, v) \in (X \times Y)^+$, where $u = x_1 \dots x_n$, $v = y_1 \dots y_n$, for $n \in \mathbb{N}$, $x_1, \dots, x_n \in X$, $y_1, \dots, y_n \in Y$. Consider any $(b_0, b_1, \dots, b_n) \in B^{n+1}$ and the product

$$\begin{aligned} &\sigma^B(b_0) \cdot \omega_{x_1, y_1}^B(b_0) \cdot \delta_{x_1}^B(b_0, b_1) \cdot \omega_{x_2, y_2}^B(b_1) \cdot \delta_{x_2}^B(b_1, b_2) \cdot \\ &\dots \cdot \omega_{x_n, y_n}^B(b_{n-1}) \cdot \delta_{x_n}^B(b_{n-1}, b_n). \end{aligned} \quad (53)$$

If for each $i \in \{1, \dots, n\}$ we have that

$$b_{i-1} = (a_{i-1}, x_i, y_i), \text{ for some } a_{i-1} \in A, \quad (54)$$

and if

$$b_n = (a_n, x, y), \text{ for some } a_n \in A, x \in X \text{ and } y \in Y, \quad (55)$$

then

$$\begin{aligned} \sigma^B(b_0) &= p \sigma^A(a_0), \quad \omega_{x_i, y_i}^B(b_{i-1}) = 1, \\ \delta_{x_i}^B(b_{i-1}, b_i) &= \mu_{x_i, y_i}^A(a_{i-1}, a_i), \end{aligned}$$

for each $i \in \{1, \dots, n\}$, and the product (53) becomes

$$\sigma^A(a_0) \cdot \mu_{x_1, y_1}^A(a_0, a_1) \cdot \mu_{x_2, y_2}^A(a_1, a_2) \cdot \dots \cdot \mu_{x_n, y_n}^A(a_{n-1}, a_n). \quad (56)$$

Otherwise, if there is $i \in \{1, \dots, n\}$ such that (54) does not hold, then $\omega_{x_i, y_i}^B(b_{i-1}) = 0$, and the whole product (53) is equal to 0.

For any $a_n \in A$ there are k elements $b_n \in B$ satisfying (55), and thus, for any $(n+1)$ -tuple $(a_0, \dots, a_n) \in A^{n+1}$ there are $k(n+1)$ -tuples $(b_0, \dots, b_n) \in B^{n+1}$ which satisfy (54) and (55). Consequently,

$$\begin{aligned} \llbracket \mathcal{B} \rrbracket_s(u, v) &= \sum_{(b_0, b_1, \dots, b_n) \in B^{n+1}} \sigma^B(b_0) \cdot \omega_{x_1, y_1}^B(b_0) \cdot \delta_{x_1}^B(b_0, b_1) \cdot \\ &\quad \cdot \omega_{x_2, y_2}^B(b_1) \cdot \delta_{x_2}^B(b_1, b_2) \cdot \dots \cdot \omega_{x_n, y_n}^B(b_{n-1}) \cdot \delta_{x_n}^B(b_{n-1}, b_n) \\ &= \sum_{(a_0, a_1, \dots, a_n) \in A^{n+1}} [p \sigma^A(a_0)] \cdot \mu_{x_1, y_1}^A(a_0, a_1) \cdot \\ &\quad \cdot \mu_{x_2, y_2}^A(a_1, a_2) \cdot \dots \cdot [k \mu_{x_n, y_n}^A(a_{n-1}, a_n)] \\ &= \sum_{(a_0, a_1, \dots, a_n) \in A^{n+1}} \sigma^A(a_0) \cdot \mu_{x_1, y_1}^A(a_0, a_1) \cdot \\ &\quad \cdot \mu_{x_2, y_2}^A(a_1, a_2) \cdot \dots \cdot [(pk) \mu_{x_n, y_n}^A(a_{n-1}, a_n)] \\ &= \sum_{(a_0, a_1, \dots, a_n) \in A^{n+1}} \sigma^A(a_0) \cdot \mu_{x_1, y_1}^A(a_0, a_1) \cdot \\ &\quad \cdot \mu_{x_2, y_2}^A(a_1, a_2) \cdot \dots \cdot \mu_{x_n, y_n}^A(a_{n-1}, a_n) \\ &= \llbracket \mathcal{A} \rrbracket(u, v), \end{aligned}$$

and hence, $\llbracket \mathcal{B} \rrbracket_s = \llbracket \mathcal{A} \rrbracket$. Clearly, $|\mathcal{B}| \leq |\mathcal{A}| \cdot |X| \cdot |Y|$. \square

7 Crisp-deterministic weighted finite automata with output

Let $\mathcal{A} = (A, X, Y, \sigma, \delta, \omega)$ be a Mealy-type weighted finite automaton over a semiring S . The weighted transition function δ is called *crisp-deterministic* if for all $x \in X$ and $a \in A$ there exists $a' \in A$ such that $\delta_x(a, a') = 1$, and $\delta_x(a, b) = 0$, for all $b \in A \setminus \{a'\}$. Also, the initial weight vector σ is *crisp-deterministic* if there exists $a_0 \in A$ such that $\sigma(a_0) = 1$, and $\sigma(a) = 0$ for all $a \in A \setminus \{a_0\}$. If both σ and δ are crisp-deterministic, then \mathcal{A} is called a *crisp-deterministic Mealy-type weighted automaton*.

Equivalently, we define a crisp-deterministic Mealy-type weighted automaton over a semiring S as a tuple $\mathcal{A} = (A, X, Y, a_0, \delta, \omega)$, where A is a non-empty set of states, $a_0 \in A$ is an initial state, $\delta : A \times X \rightarrow A$ is a transition function and $\omega : A \times X \times Y \rightarrow S$ is a weighted output function. For any $x \in X$ we define $\delta_x : A \rightarrow A$ by $\delta_x(a) = \delta(a, x)$, for all $a \in A$, and for any $u \in X^*$ we define the transition function $\delta_u : A \rightarrow A$ as follows: For any $a \in A$ we set $\delta_\varepsilon(a) = a$, and for $a \in A$, $u \in X^*$ and $x \in X$, we set $\delta_{ux}(a) = \delta_x(\delta_u(a))$.

A *crisp-deterministic Moore-type weighted automaton* over S is defined as a tuple $\mathcal{A} = (A, X, Y, a_0, \delta, \omega)$, where everything is the same as in the definition of a crisp-deterministic Mealy-type weighted automaton except

the weighted output function, for which we assume that $\omega : A \times Y \rightarrow S$.

Given that crisp-deterministic Mealy-type weighted automata are a special type of the general Mealy-type weighted automata, the $1n$ -semantics, $n1$ -semantics and sequential semantics for these automata are those that are defined in Section 4. Similarly, the definitions of $1n$ -semantics and $n1$ -semantics for Moore-type weighted automata given in Section 5 apply also to crisp-deterministic Moore-type weighted automata. However, in the case of crisp-deterministic Mealy-type and Moore-type weighted automata it is natural to consider the following semantics for which we prove that they are equivalent to all the above listed semantics.

Definition 7.1 (Crisp-deterministic semantics) The *cd-behavior* of a crisp-deterministic Mealy-type weighted automaton $\mathcal{A} = (A, X, Y, a_0, \delta, \omega)$ is the function $\llbracket \mathcal{A} \rrbracket_{cd} : (X \times Y)^* \rightarrow S$ defined by

$$\llbracket \mathcal{A} \rrbracket_{cd}(\varepsilon, \varepsilon) = 1, \quad (57)$$

and

$$\begin{aligned} \llbracket \mathcal{A} \rrbracket_{cd}(u, v) = \\ = \omega_{x_1, y_1}(a_0) \cdot \omega_{x_2, y_2}(\delta_{x_1}(a_0)) \cdot \dots \cdot \omega_{x_n, y_n}(\delta_{x_1 \dots x_{n-1}}(a_0)), \end{aligned} \quad (58)$$

for all $u = x_1 x_2 \dots x_n \in X^*$ and $v = y_1 y_2 \dots y_n \in Y^*$.

Similarly, by the *cd-behavior* of a crisp-deterministic Moore-type weighted automaton $\mathcal{A} = (A, X, Y, a_0, \delta, \omega)$ we mean the function $\llbracket \mathcal{A} \rrbracket_{cd} : (X \times Y)^* \rightarrow S$ defined by

$$\llbracket \mathcal{A} \rrbracket_{cd}(\varepsilon, \varepsilon) = 1, \quad (59)$$

and

$$\begin{aligned} \llbracket \mathcal{A} \rrbracket_{cd}(u, v) = \\ = \omega_{y_1}(\delta_{x_1}(a_0)) \cdot \omega_{y_2}(\delta_{x_1 x_2}(a_0)) \cdot \dots \cdot \omega_{y_n}(\delta_{x_1 \dots x_n}(a_0)), \end{aligned} \quad (60)$$

for all $u = x_1 x_2 \dots x_n \in X^*$ and $v = y_1 y_2 \dots y_n \in Y^*$.

Now we show that for all crisp-deterministic Mealy-type and Moore-type weighted automata the above defined semantics coincide with all semantics defined in Sections 4 and 5 for the general Mealy-type and Moore-type weighted automata.

Theorem 7.1 *If $\mathcal{A} = (A, X, Y, a_0, \delta, \omega)$ is a crisp-deterministic Mealy-type weighted automaton then*

$$\llbracket \mathcal{A} \rrbracket_{cd} = \llbracket \mathcal{A} \rrbracket_{1n} = \llbracket \mathcal{A} \rrbracket_{n1} = \llbracket \mathcal{A} \rrbracket_s, \quad (61)$$

and if \mathcal{A} is a crisp-deterministic Moore-type weighted automaton then

$$\llbracket \mathcal{A} \rrbracket_{cd} = \llbracket \mathcal{A} \rrbracket_{1n} = \llbracket \mathcal{A} \rrbracket_{n1}. \quad (62)$$

Proof If \mathcal{A} is a crisp-deterministic Mealy-type weighted automaton, it is easy to check that the rightmost terms in equations (16), (21) and (23) become equal to 1, while the rightmost terms in equations (17), (22) and (24) are converted into the term on the right-hand side of equation (58). Therefore, (61) holds.

Similarly, if \mathcal{A} is a crisp-deterministic Moore-type weighted automaton, then the rightmost terms in equations (31) and (36) are equal to 1, whereas the rightmost terms in (37) and (37) are transformed into the term on the right-hand side of (60). Thus, we conclude that (62) is true. \square

References

1. Cheng W, Mo ZW (2004), Minimization algorithm of fuzzy finite automata. *Fuzzy Sets and Systems* 141: 439–448.
2. Lei HX, Li YM (2007) Minimization of states in automata theory based on finite lattice-ordered monoids. *Information Sciences* 177: 1413–1421.
3. Li YM, Pedrycz W (2006) The equivalence between fuzzy Mealy and fuzzy Moore machines. *Soft Computing* 10: 953–959.
4. Liu J, Moa ZW, Qiu D, Wang Y (2009) Products of Mealy-type fuzzy finite state machines. *Fuzzy Sets and Systems* 160: 2401–2415.
5. Mohri M (2009) Weighted automata algorithms. In: Droste M, Kuich W, Vogler H (eds.), *Handbook of Weighted Automata*, Monographs in Theoretical Computer Science, An EATCS Series, Springer-Verlag, pp. 213–254.
6. Mordeson JN, Nair PS (1996) Fuzzy Mealy machines. *Kybernetes* 25(3): 18–33.
7. Peeva K (1988) Behaviour, reduction and minimization of finite L-automata. *Fuzzy Sets and Systems* 28: 171–181.
8. Peeva K (2004) Finite L-fuzzy machines. *Fuzzy Sets and Systems* 141: 415–437.
9. Peeva K, Kyosev Y (2004) *Fuzzy Relational Calculus: Theory, Applications, and Software*. World Scientific, Singapore.
10. Peeva K, Zahariev Z (2008) Computing behavior of finite fuzzy machines – Algorithm and its application to reduction and minimization. *Information Sciences* 178: 4152–4165.
11. Petković T (2006) Congruences and homomorphisms of fuzzy automata. *Fuzzy Sets and Systems* 157: 444–458.
12. Wu LH, Qiu DW (2010) Automata theory based on complete residuated lattice-valued logic: Reduction and minimization. *Fuzzy Sets and Systems* 161: 1635–1656.
13. Xing H, Qiu DW, Liu FC, Fan ZJ (2007) Equivalence in automata theory based on complete residuated lattice-valued logic. *Fuzzy Sets and Systems* 158: 1407–1422.

